IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of Mark E. Valenti

Serial No.: 09/613,387 Group Art Unit: 2755

Filed: 07/11/2000 Examiner:

For: **SYSTEM AND METHOD FOR INTERNET BROADCAST SEARCHING**

Enclosed please find the following:

1. Petition to Make Special;
2. Fee in the amount of $130.00 in pursuant to 37 CFR 1.17(h)
3. Certificate of Express mailing.

The Commissioner is hereby authorized to charge any fee deficiency, or credit any overpayment, to Deposit Account No. 18-1579. The Commissioner is also authorized to charge Deposit Account No. 18-1579 for any future fees connected in any way to this application. Two copies of this letter are enclosed.

Respectfully submitted,

Jon L. Roberts, Esq.
Registration No. 31,293
Roberts Abokhair & Mardula, LLC
11800 Sunrise Valley Drive, Suite 1000
Reston, VA 20191-5302
(703) 391-2900

#6

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of Mark E. Valenti

Serial No.: 09/613,387 Group Art Unit: 2755

Filed: 07/11/2000 Examiner:

For: **SYSTEM AND METHOD FOR INTERNET BROADCAST SEARCHING**

Commissioner for Patents
Box PATENT APPLICATION
Washington, D.C. 20231

PETITION TO MAKE SPECIAL AND

PETITION FOR ACCELERATED EXAMINATION PURSUANT TO MPEP

708.02 VII (C)

Dear Sir:

Applicant hereby requests that an accelerated examination of the above referenced patent application be made based upon MPEP 708.02 VII (C).

The fee pursuant to 37 CFR 1.17(h) is included herewith.

Background:

On January 12, 2001, a patent application under the provisions of the Patent Cooperation treaty was filed and received serial No. PCT/US01/01096.

A preexamination search was made and incorporated into the IPER.

This preexamination search was made by the USPTO acting as the receiving office and as the office designated as the International Search Authority (ISA).

On November 6, 2001 a combined search and International Preliminary examination Report was issued (Exhibit A). That IPER cited certain references which are

attached hereto as Exhibit B are deem the references most closely related to the pending application.

An Information Disclosure Statement was filed on 12/14/00 citing certain references which were discussed in the application. Copies of those references were provided to the USPTO as part of the submission of the IDS.

Following is a discussion of the references cited by the ISA and together with the references discussed in the pending application constitute the detailed discussion called fro in MPEP 708.02 VII (E)

USP 6,236,991 to Frauenhoder et al.

USP No. USP 6,236,991 to Frauenhoder et al. was issued on May 22, 2001 for a "Method and System for Providing Access for Categorized Information from Online Internet and Intranet Sources." The broadest claim of the patent is claims one which is set forth below with a discussion of that claim.

1. A method of searching a network of interconnected computers and servers comprising:

categorizing information stored on a plurality of information servers connected to a network to form categorization information;

collecting and storing the categorization information and network addresses of the information servers on a plurality of IBSP servers;

transmitting the categorization information and network addresses of the plurality of information servers from an IBSP server to user nodes, broadcast server nodes, or firewall server nodes over the network;

accepting a query on a user node connected to the network;

transmitting the query from the user node directly to a plurality of information servers or to a broadcast server or a firewall server over the network;

the broadcast server or firewall server receiving and transmitting the user node query to the plurality of information servers;

the information servers searching themselves for information responsive to the user node query; and

each of the plurality of information servers transmitting information responsive to the user node query to the user node or the firewall server for forwarding to the user node when responsive information if found.

09/613,387	USP 6,236,991
A method of searching a network of interconnected computers and servers comprising:	A system for allowing access to electronic documents
Categorizing information stored on a plurality of information servers connected to a network to form categorization information;	Categories of information on the server are not determined. Once documents are retrieved, then topics are derived from the retrieved documents.
collecting and storing the categorization information and network addresses of the information servers on a plurality of IBSP servers;	'991 gathers documents rather than categorization information. Topics are only derived after the documents themselves are retrieved and stored to a provider computer.
transmitting the categorization information and network addresses of the plurality of information servers from an IBSP server to user nodes, broadcast server nodes, or firewall server nodes over the network;	'991 transmits first and second electronic documents themselves to a "provider computer." The topic information is not derived until the documents are stored on the provider computer.
accepting a query on a user node connected to the network;	Obtains a user profile of interests rather than a specific query for information. The query is determined from the user profile.
transmitting the query from the user node directly to a plurality of information servers or to a broadcast server or a firewall server over the network;	The query is transmitted to web sites in order to retrieve documents that match the query rather that categorization information
the broadcast server or firewall server receiving and transmitting the user node query to the plurality of information servers;	The user profile of '991 is transmitted as a query after analysis.
the information servers searching themselves for information responsive to the user node query; and	Documents responsive to the query are retrieved from the "electronic content source."
each of the plurality of information servers transmitting information responsive to the user node query to the user node or the firewall server for forwarding to the user node when responsive information if found.	Documents retrieved are stored on a provider computer for subsequent access by a user. The documents are not returned to the user from the content source but from the provider computer.

As can bee seen from the above comparison, the '991 patent is designed to review a user profile and to convert that user profile into a query for retrieval of document without the user requesting them. When the user want a document of a type that is in the user's profile, retrieval of the document can be more efficient since it has been stored at a provider computer closer to the user.

In contrast, the present invention causes a categorization of the data that is on a variety of content servers. The categories and address of the content providers is stored, not the documents themselves. When a user want a query to be processed, a request is made and servers that have the appropriate category of data and addresses, direct the query more efficiently to the content server that has the requisite data. User profiles are not involved in the process of the present invention.

USP 6,195,654 to Wachtel

The 654 application does not relate directly to the present invention. This patent is drawn to improving search results by viewing the behavior of a internet user, analyzing that behavior and returning search results that more closely match the behavior of the searcher.

The 654 patent comprises a client database where interest categories, information accessed, profile characteristics and rejected categories of data are stored concerning each client. This information is continually updated. When a search is then requested, comparison to previous searches relating to the client is performed so that data that more closely matches the clients interest can be returned.

In contrast, the present invention deals with categorization of information that is on information servers in a wide variety of locations and sending the categories of information and the addresses of that information to servers at various locations on the network. When a user desires information, the request is sent to certain servers that look for the categories of information the user desires and sends the request only to those places where the information is present. Thus the search efficiency is improved.

USP No. 6,085,225 to Nakajima et al.

USP No. 6,085,225 to Nakajima et al. was issued for an "Information System, Information Storing and Providing Apparatus, and Information Receiving Apparatus." Nakajima et al. is not drawn to a search model, but instead is drawn to a model for updating information based on newer information being available from a source. As such, it lacks the 1) collecting and storing of categorization information and network addresses of the information servers on a plurality of IBSP servers; 2) transmitting queries from user nodes directly to a plurality of information servers or to a broadcast server or a firewall server over the network; 3) the broadcast server or firewall server receiving and transmitting the user node query to the plurality of information servers; and 4) each of the plurality of information servers transmitting information responsive to user node queries to user nodes or firewall servers for forwarding to user nodes when responsive information if found. A comparison with claim 1 (the broadest claim) is presented below:

A method of searching a network of interconnected computers and servers comprising:	Nakajima et al. discloses a network with "an information storing and providing apparatus" and "a plurality of information receipt apparatuses" - i.e., a single server for plural clients
Categorizing information stored on a plurality of information servers connected to a network to form categorization information;	In one embodiment of Nakajima et al., "information storing and providing apparatus (server) comprises means for generating first category information denoting a category of the supply information;"
collecting and storing the categorization information and network addresses of the information servers on a plurality of IBSP servers;	Not taught or fairly suggested by Nakajima et al.
transmitting the categorization information and network addresses of the plurality of information servers from an IBSP server to user nodes, broadcast server nodes, or firewall server nodes over the network;	The "information providing apparatus" (server) of Nakajima et al. includes "means for broadcasting the category information via the network" but has no need for network addresses since the information providing apparatus is the sole information source
accepting a query on a user node connected to the network;	Nakajima et al. does not disclose accepting a query on a user node, but does disclose "the plurality of the information apparatuses (clients) each comprise means for storing second category information with respect to supply information necessary therefor"
transmitting the query from the user node directly to a plurality of information servers or to a broadcast server or a firewall server over the network;	Not taught or fairly suggested by Nakajima et al.
the broadcast server or firewall server receiving and transmitting the user node query to the plurality of information servers;	Not taught or fairly suggested by Nakajima et al.
the information servers searching themselves for information responsive to the user node query; and	Not taught or fairly suggested by Nakajima et al.; instead, the comparison is done at the information apparatus (client) with "means for receiving the first category information;

	means for comparing the first category information with the second category information to determine whether or not the supply information from the information storing and providing apparatus is necessary"
each of the plurality of information servers transmitting information responsive to the user node query to the user node or the firewall server for forwarding to the user node when responsive information if found.	Not taught or fairly suggested by Nakajima et al.; instead, the information apparatus (client) includes "means for presenting the supply information from the information storing and providing apparatus upon determination that the supply information is necessary"

In view of the above detailed explanations and those contained in the application for the present invention, it is respectfully submitted that the claimed subject matter is patentable distinct from the above referenced patents and those cited in the application.

If there are any questions concerning this petition to Make Special, please contact the undersigned.

Respectfully Submitted,

Jon L. Roberts
Registration No.: 31,293
Roberts, Abokhair, & Mardula, LLC
11800 Sunrise Valley Drive
Suite 1000
Reston, VA 20191

JLR/vjm

Exhibit A

PATENT COOPERATION TREATY

From the
INTERNATIONAL PRELIMINARY EXAMINING AUTHORITY

To: JON L. ROBERTS
ROBERTS ABOKHAIR AND MARDULA LLC
11800 SUNRISE VALLEY DRIVE, SUITE 1000
RESTON, VA 20191

PCT

NOTIFICATION OF TRANSMITTAL OF INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Rule 71.1)

Date of Mailing *(day/month/year)*	06 NOV 2001
Applicant's or agent's file reference	IMPORTANT NOTIFICATION
2639-001PCT2	

International application No.	International filing date *(day/month/year)*	Priority Date *(day/month/year)*
PCT/US01/01096	12 JANUARY 2001	13 JANUARY 2000

Applicant
INTERLINK NETWORK RESOURCES, INC.

1. The applicant is hereby notified that this International Preliminary Examining Authority transmits herewith the international preliminary examination report and its annexes, if any, established on the international application.

2. A copy of the report and its annexes, if any, is being transmitted to the International Bureau for communication to all the elected Offices.

3. Where required by any of the elected Offices, the International Bureau will prepare an English translation of the report (but not of any annexes) and will transmit such translation to those Offices.

4. **REMINDER**

The applicant must enter the national phase before each elected Office by performing certain acts (filing translations and paying national fees) within 30 months from the priority date (or later in some Offices)(Article 39(1))(see also the reminder sent by the International Bureau with Form PCT/IB/301).

Where a translation of the international application must be furnished to an elected Office, that translation must contain a translation of any annexes to the international preliminary examination report. It is the applicant's responsibility to prepare and furnish such translation directly to each elected Office concerned.

For further details on the applicable time limits and requirements of the elected Offices, see Volume II of the PCT Applicant's Guide.

Name and mailing address of the IPEA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	AYAZ R. SHEIKH
Facsimile No. (703) 305-3230	Telephone No. (703) 305-9648

Form PCT/IPEA/416 (July 1992) *

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference 2639-001PCT2	FOR FURTHER ACTION	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
International application No. PCT/US01/01096	International filing date *(day/month/year)* 12 JANUARY 2001	Priority date *(day/month/year)* 13 JANUARY 2000
International Patent Classification (IPC) or national classification and IPC IPC(7): G06f 15/16 and US Cl.: 709/219		
Applicant INTERLINK NETWORK RESOURCES, INC.		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 4 sheets.

   ☐ This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority. (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

   These annexes consist of a total of _____ sheets.

3. This report contains indications relating to the following items:

   I ☒ Basis of the report

   II ☐ Priority

   III ☐ Non-establishment of report with regard to novelty, inventive step or industrial applicability

   IV ☐ Lack of unity of invention

   V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

   VI ☒ Certain documents cited

   VII ☐ Certain defects in the international application

   VIII ☐ Certain observations on the international application

Date of submission of the demand	Date of completion of this report
09 AUGUST 2001	13 SEPTEMBER 2001
Name and mailing address of the IPEA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer Peggy Harrod AYAZ R. SHEIKH Telephone No. (703) 305-9648

Form PCT/IPEA/409 (cover sheet) (July 1998) ★

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

rnational application No.
PCT/US01/01096

I. Basis of the report

1. With regard to the **elements** of the international application:*

[X] the international application as originally filed

[X] the description:
pages 1-13 , as originally filed
pages NONE , filed with the demand
pages NONE , filed with the letter of _____

[X] the claims:
pages 14-17 , as originally filed
pages NONE , as amended (together with any statement) under Article 19
pages NONE , filed with the demand
pages NONE , filed with the letter of _____

[X] the drawings:
pages 1-3 , as originally filed
pages NONE , filed with the demand
pages NONE , filed with the letter of _____

[X] the sequence listing part of the description:
pages NONE , as originally filed
pages NONE , filed with the demand
pages NONE , filed with the letter of _____

2. With regard to the **language**, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.
These elements were available or furnished to this Authority in the following language _____ which is:

[ ] the language of a translation furnished for the purposes of international search (under Rule 23.1(b)).

[ ] the language of publication of the international application (under Rule 48.3(b)).

[ ] the language of the translation furnished for the purposes of international preliminary examination (under Rules 55.2 and/or 55.3).

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:

[ ] contained in the international application in printed form.

[ ] filed together with the international application in computer readable form.

[ ] furnished subsequently to this Authority in written form.

[ ] furnished subsequently to this Authority in computer readable form.

[ ] The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.

[ ] The statement that the information recorded in computer readable form is identical to the writen sequence listing has been furnished.

4. [X] The amendments have resulted in the cancellation of:

[X] the description, pages NONE

[X] the claims, Nos. NONE

[X] the drawings, sheets/~~fig~~ NONE

5. [ ] This report has been drawn as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)).**

** Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report since they do not contain amendments (Rules 70.16 and 70.17).*

***Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report.*

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. statement

Novelty (N)	Claims	1-11	YES
	Claims	NONE	NO
Inventive Step (IS)	Claims	1-11	YES
	Claims	NONE	NO
Industrial Applicability (IA)	Claims	1-11	YES
	Claims	NONE	NO

2. citations and explanations (Rule 70.7)

Claims 1-11 meet the criteria set out in PCT Article 33(2)-(4), because the prior art does not teach or fairly suggest using an Internet Broadcast Search Paradigm (IBSP) to speed up searching process on the Internet wherein categorizing information and network addresses are broadcasted from the IBSP and stored at multi-levels of computers on the network such as users, broadcast servers and firewalls.

---------- NEW CITATIONS ----------

NONE

VI. Certain documents cited

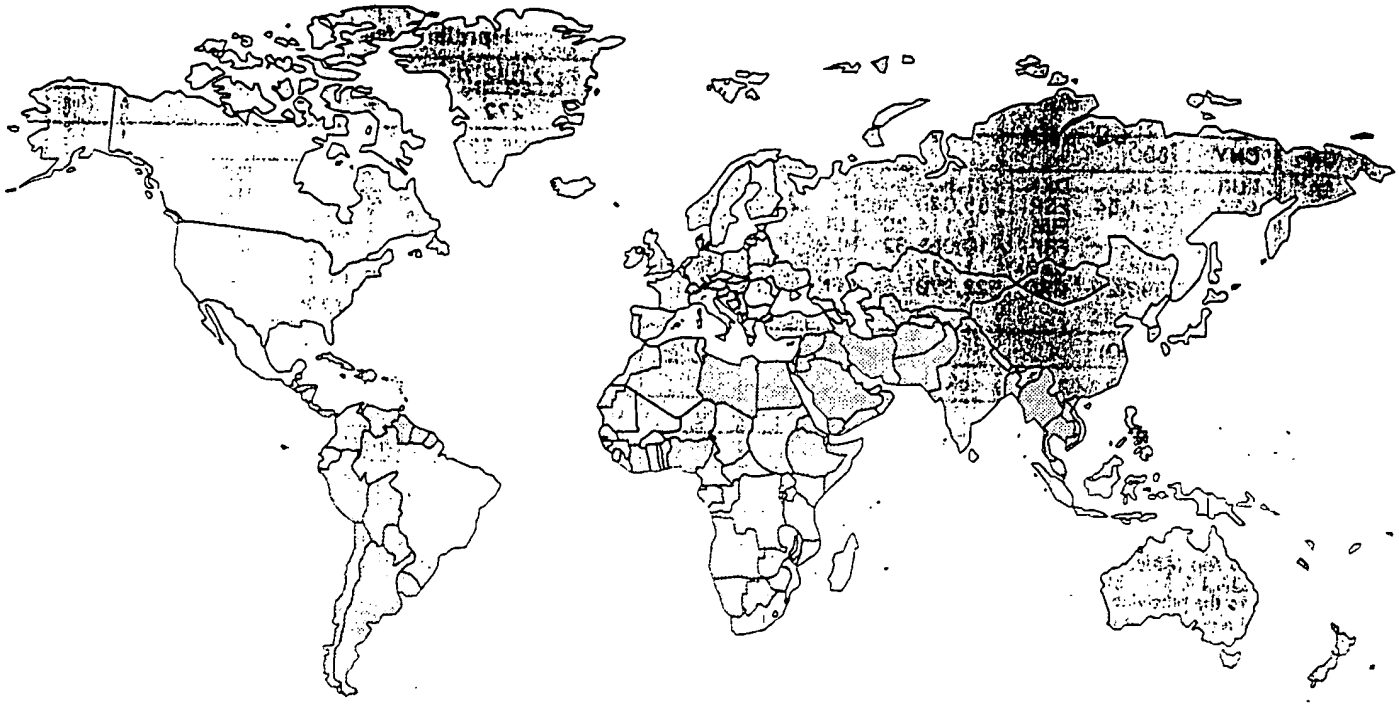
1. Certain published documents (Rule 70.10)

Application No. Patent No.	Publication Date *(day/month/year)*	Filing Date *(day/month/year)*	Priority date (valid claim) *(day/month/year)*
US, B1, 6,236,991	22 MAY 2001	26 NOVEMBER 1997	NONE
US, B1, 6,195,654	27 FEBRUARY 2001	18 NOVEMBER 1996	NONE
US, A, 6,085,225	04 JULY 2000	27 AUGUST 1997	28 AUGUST 1996

2. Non-written disclosures (Rule 70.9)

Kind of non-written disclosure	Date of non-written disclosure *(day/month/year)*	Date of written disclosure referring to non-written disclosure *(day/month/year)*

PCT CONTRACTING STATES AND TWO-LETTER CODES (115 on 1 October 2001)



AE United Arab Emirates
AG Antigua and Barbuda
AL Albania[1]
AM Armenia (EA)
AT Austria (EP)
AU Australia
AZ Azerbaijan (EA)
BA Bosnia and Herzegovina
BB Barbados
BE Belgium (EP)[2]
BF Burkina Faso (OA)[2]
BG Bulgaria
BJ Benin (OA)[2]
BR Brazil
BY Belarus (EA)
BZ Belize
CA Canada
CF Central African Republic (OA)[2]
CG Congo (OA)[2]
CH Switzerland (EP)
CI Côte d'Ivoire (OA)[2]
CM Cameroon (OA)[2]
CN China
CO Colombia
CR Costa Rica
CU Cuba
CY Cyprus (EP)[2]
CZ Czech Republic
DE Germany (EP)
DK Denmark (EP)
DM Dominica
DZ Algeria
EC Ecuador
EE Estonia
ES Spain (EP)
FI Finland (EP)
FR France (EP)[2]
GA Gabon (OA)[2]
GB United Kingdom (EP)
GD Grenada
GE Georgia
GH Ghana (AP)
GM Gambia (AP)
GN Guinea (OA)[2]
GQ Equatorial Guinea (OA)[2]
GR Greece (EP)[2]
GW Guinea-Bissau (OA)[2]
HR Croatia
HU Hungary
ID Indonesia
IE Ireland (EP)[2]
IL Israel
IN India
IS Iceland
IT Italy (EP)[2]
JP Japan
KE Kenya (AP)
KG Kyrgyzstan (EA)
KP Democratic People's Republic of Korea
KR Republic of Korea
KZ Kazakhstan (EA)
LC Saint Lucia
LI Liechtenstein (EP)
LK Sri Lanka
LR Liberia
LS Lesotho (AP)
LT Lithuania[1]
LU Luxembourg (EP)
LV Latvia[1]
MA Morocco
MC Monaco (EP)[2]
MD Republic of Moldova (EA)
MG Madagascar
MK The former Yugoslav Republic of Macedonia[1]
ML Mali (OA)[2]
MN Mongolia
MR Mauritania (OA)[2]
MW Malawi (AP)
MX Mexico
MZ Mozambique (AP)
NE Niger (OA)[2]
NL Netherlands (EP)[2]
NO Norway
NZ New Zealand
OM Oman (from 26 October 2001)
PH Philippines
PL Poland
PT Portugal (EP)
RO Romania[1]
RU Russian Federation (EA)
SD Sudan (AP)
SE Sweden (EP)
SG Singapore
SI Slovenia[1]
SK Slovakia
SL Sierra Leone (AP)
SN Senegal (OA)[2]
SZ Swaziland (AP)[2]
TD Chad (OA)[2]
TG Togo (OA)[2]
TJ Tajikistan (EA)
TM Turkmenistan (EA)
TN Tunisia (from 10 December 2001)
TR Turkey (EP)
TT Trinidad and Tobago
TZ United Republic of Tanzania (AP)
UA Ukraine
UG Uganda (AP)
US United States of America
UZ Uzbekistan
VN Viet Nam
YU Yugoslavia
ZA South Africa
ZM Zambia (from 15 November 2001)
ZW Zimbabwe (AP)

1 Extension of European patent possible.
2 May only be designated for a regional patent (the "national route" via the PCT has been closed).

Where a State can be designated for a regional patent, the two-letter code for the regional patent concerned is indicated in parentheses (AP = ARIPO patent, EA = Eurasian patent, EP = European patent, OA = OAPI patent).

Important:

This list includes all States that have adhered to the PCT by the date shown in the heading. Any State indicated in ***bold italics*** has adhered to the PCT but was not yet bound by the PCT on the date of issue of the latest version of the request form. Where any State has adhered to but is not yet bound by the PCT, the date on which it will become bound is shown in parentheses; it cannot be designated in international applications filed before that date. If the applicant wishes to designate, for the purposes of a national patent, any State which is bound by the PCT on the date on which the international application is filed but which is not listed in the request form, he must add it in Box No. V of the request form and mark the corresponding check-box.

Applicants should always use the latest versions of the request form (PCT/RO/101) and demand form (PCT/IPEA/401), updated versions of which are normally issued twice yearly. The latest versions are dated July 2001, and can be printed from the WIPO Internet site, in editable PDF format, at: http://www.wipo.int/pct/en/forms/index.htm, or obtained from receiving Offices or the International Bureau, or, in the case of the demand form, also from International Preliminary Examining Authorities.

Box No.V DESIGNATION OF STATES *Mark the applicable check-boxes below; at least one must be marked.*

The following designations are hereby made under Rule 4.9(a):

Regional Patent

☒ **AP ARIPO Patent:** GH Ghana, GM Gambia, KE Kenya, LS Lesotho, MW Malawi, MZ Mozambique, SD Sudan, SL Sierra Leone, SZ Swaziland, TZ United Republic of Tanzania, UG Uganda, ZW Zimbabwe, and any other State which is a Contracting State of the Harare Protocol and of the PCT

☒ **EA Eurasian Patent:** AM Armenia, AZ Azerbaijan, BY Belarus, KG Kyrgyzstan, KZ Kazakhstan, MD Republic of Moldova, RU Russian Federation, TJ Tajikistan, TM Turkmenistan, and any other State which is a Contracting State of the Eurasian

☒ **EP European Patent:** AT Austria, BE Belgium, CH & LI Switzerland and Liechtenstein, CY Cyprus, DE Germany, DK Denmark, ES Spain, FI Finland, FR France, GB United Kingdom, GR Greece, IE Ireland, IT Italy, LU Luxembourg, MC Monaco, NL Netherlands, PT Portugal, SE Sweden, TR Turkey, and any other State which is a Contracting State of the European Patent Convention and of the PCT

☒ **OA OAPI Patent:** BF Burkina Faso, BJ Benin, CF Central African Republic, CG Congo, CI Côte d'Ivoire, CM Cameroon, GA Gabon, GN Guinea, GW Guinea-Bissau, ML Mali, MR Mauritania, NE Niger, SN Senegal, TD Chad, TG Togo, and any other State which is a member State of OAPI and a Contracting State of the PCT *(if other kind of protection or treatment desired, specify on dotted line)* .....

National Patent *(if other kind of protection or treatment desired, specify on dotted line):*

☒ AE	United Arab Emirates	☒ GH	Ghana	☒ MX	Mexico
☒ AG	Antigua and Barbuda	☒ GM	Gambia	☒ MZ	Mozambique
☒ AL	Albania	☒ HR	Croatia	☒ NO	Norway
☒ AM	Armenia	☒ HU	Hungary	☒ NZ	New Zealand
☒ AT	Austria	☒ ID	Indonesia	☒ PL	Poland
☒ AU	Australia	☒ IL	Israel	☒ PT	Portugal
☒ AZ	Azerbaijan	☒ IN	India	☒ RO	Romania
☒ BA	Bosnia and Herzegovina	☒ IS	Iceland	☒ RU	Russian Federation
☒ BB	Barbados	☒ JP	Japan		
☒ BG	Bulgaria	☒ KE	Kenya	☒ SD	Sudan
☒ BR	Brazil	☒ KG	Kyrgyzstan	☒ SE	Sweden
☒ BY	Belarus	☒ KP	Democratic People's Republic of Korea	☒ SG	Singapore
☒ BZ	Belize			☒ SI	Slovenia
☒ CA	Canada	☒ KR	Republic of Korea	☒ SK	Slovakia
☒ CH & LI	Switzerland and Liechtenstein	☒ KZ	Kazakhstan	☒ SL	Sierra Leone
☒ CN	China	☒ LC	Saint Lucia	☒ TJ	Tajikistan
☒ CO	Colombia	☒ LK	Sri Lanka	☒ TM	Turkmenistan
☒ CR	Costa Rica	☒ LR	Liberia	☒ TR	Turkey
☒ CU	Cuba	☒ LS	Lesotho	☒ TT	Trinidad and Tobago
☒ CZ	Czech Republic	☒ LT	Lithuania		
☒ DE	Germany	☒ LU	Luxembourg	☒ TZ	United Republic of Tanzania
☒ DK	Denmark	☒ LV	Latvia	☒ UA	Ukraine
☒ DM	Dominica	☒ MA	Morocco	☒ UG	Uganda
☒ DZ	Algeria	☒ MD	Republic of Moldova	☐ US	United States of America
☒ EC	Ecuador				
☒ EE	Estonia	☒ MG	Madagascar	☒ UZ	Uzbekistan
☒ ES	Spain	☒ MK	The former Yugoslav Republic of Macedonia	☒ VN	Viet Nam
☒ FI	Finland			☒ YU	Yugoslavia
☒ GB	United Kingdom	☒ MN	Mongolia	☒ ZA	South Africa
☒ GD	Grenada	☒ MW	Malawi	☒ ZW	Zimbabwe
☒ GE	Georgia				

Check-boxes below reserved for designating States which have become party to the PCT after issuance of this sheet:

☒ EC Ecuador	☐	☐
☒ OM Oman	☐	☐

Precautionary Designation Statement: In addition to the designations made above, the applicant also makes under Rule 4.9(b) all other designations which would be permitted under the PCT except any designation(s) indicated in the Supplemental Box as being excluded from the scope of this statement. The applicant declares that those additional designations are subject to confirmation and that any designation which is not confirmed before the expiration of 15 months from the priority date is to be regarded as withdrawn by the applicant at the expiration of that time limit. *(Confirmation (including fees) must reach the receiving Office within the 15-month time limit.)*

Exhibit B

US006236991B1

(12) **United States Patent**
Frauenhofer et al.

(10) **Patent No.: US 6,236,991 B1**
(45) **Date of Patent: May 22, 2001**

(54) **METHOD AND SYSTEM FOR PROVIDING ACCESS FOR CATEGORIZED INFORMATION FROM ONLINE INTERNET AND INTRANET SOURCES**

(75) Inventors: **Thomas Valentine Frauenhofer**, Stony Point, NY (US); **Joaquin Manuel Marques**, Greenwich, CT (US); **Michael Edward Moran**, Ridgewood, NJ (US); **Subhas Palchowdhury**, Mahopac, NY (US); **Jeffrey Stephen Schaffer**, Ridgefield, CT (US)

(73) Assignee: **International Business Machines Corp.**, Armonk, NY (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **08/979,861**

(22) Filed: **Nov. 26, 1997**

(51) **Int. Cl.**[7] **G06F 17/30**

(52) **U.S. Cl.** **707/6**; 707/4; 707/10; 705/10; 705/27

(58) **Field of Search** 707/6, 7, 10, 104, 707/500, 513, 515, 516, 531, 4; 395/200.57, 200.48, 200.49, 200.59, 200.3

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,659,732	*	8/1997	Kirsch	707/10
5,706,507	*	1/1998	Schloss	707/104
5,740,549	*	4/1998	Reilly et al.	705/14
5,754,939	*	5/1998	Herz et al.	707/10
5,761,662	*	6/1998	Dasan	707/10
5,774,652	*	6/1998	Smith	395/186
5,778,367	*	7/1998	Wesinger, Jr. et al.	707/10
5,790,664	*	8/1998	Coley et al.	380/4
5,815,665	*	9/1998	Teper et al.	395/200.59
5,819,284	*	10/1998	Farber et al.	707/104
5,826,014	*	10/1998	Coley et al.	395/187.01
5,835,087	*	11/1998	Herz et al.	345/327
5,835,905	*	11/1998	Pirolli et al.	707/3
5,862,339	*	1/1999	Bonnaure et al.	395/200.57
5,870,550	*	2/1999	Wesinger, Jr. et al.	395/200.48
5,881,234	*	3/1999	Schwob	395/200.49
5,887,133	*	3/1999	Brown et al.	395/200.3
5,895,470	*	4/1999	Pirolli et al.	707/102
5,924,090	*	7/1999	Krellenstein	707/5
5,931,907	*	8/1999	Davies et al.	709/218
5,943,670	*	8/1999	Prager	707/5
5,987,454	*	11/1999	Hobbs	707/4

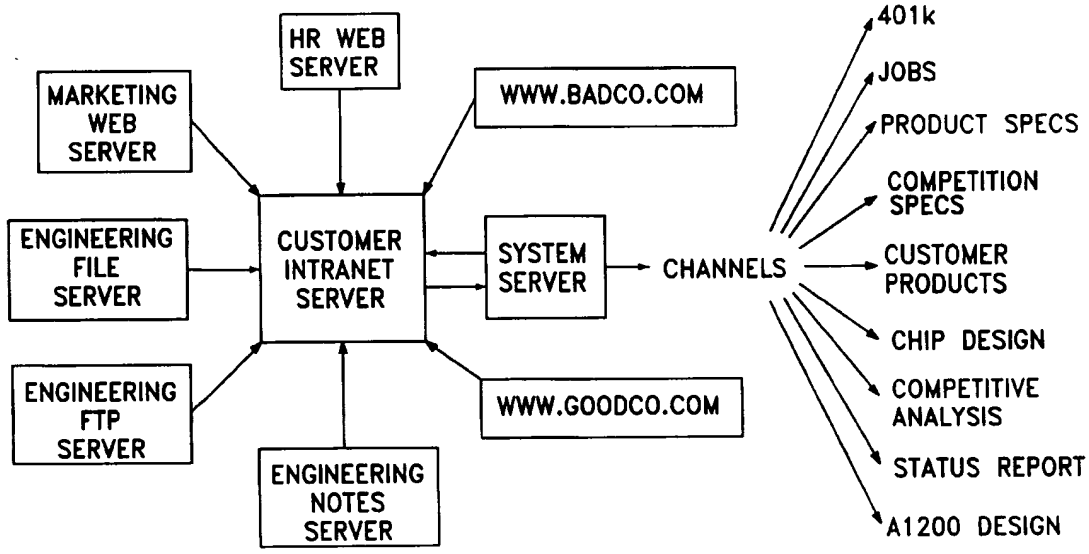
* cited by examiner

Primary Examiner—Thomas Black
Assistant Examiner—Frantz Coby
(74) *Attorney, Agent, or Firm*—David M. Shofi; Anne Vachon Dougherty

(57) **ABSTRACT**

A system for collecting, categorizing and searching meta-data about content provided on the internet and/or intranet for delivery in accordance with customized user profiles. The system collects internet information and categorizes same for provision at a customer's intranet server. The system is additionally adapted to either passively receive or actively collect and categorize internally-provided content for delivery with the externally gathered and categorized content and for matching to user profiles.

22 Claims, 2 Drawing Sheets



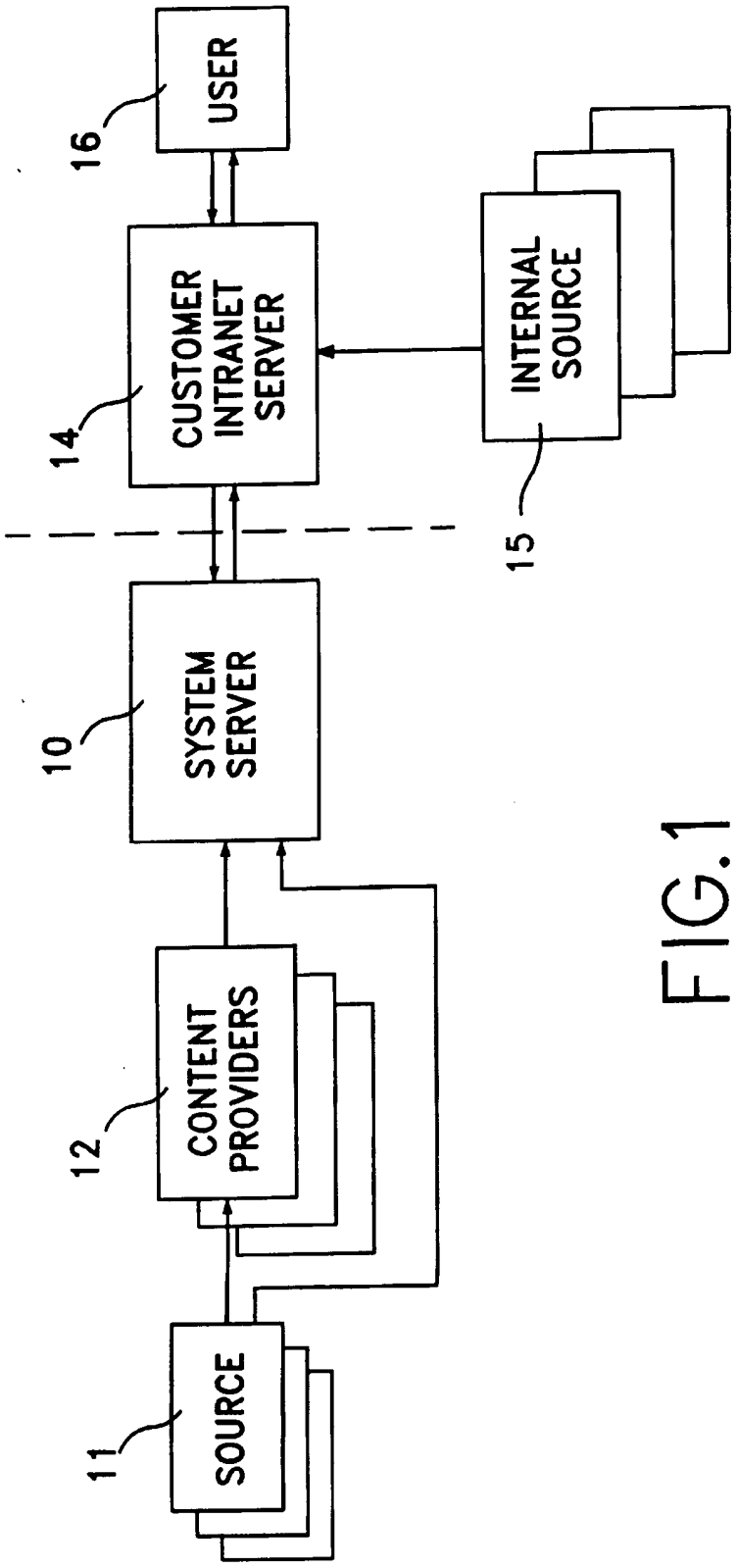

USER
16
CUSTOMER INTRANET SERVER
14
INTERNAL SOURCE
15
SYSTEM SERVER
10
CONTENT PROVIDERS
12
SOURCE
11

FIG. 1

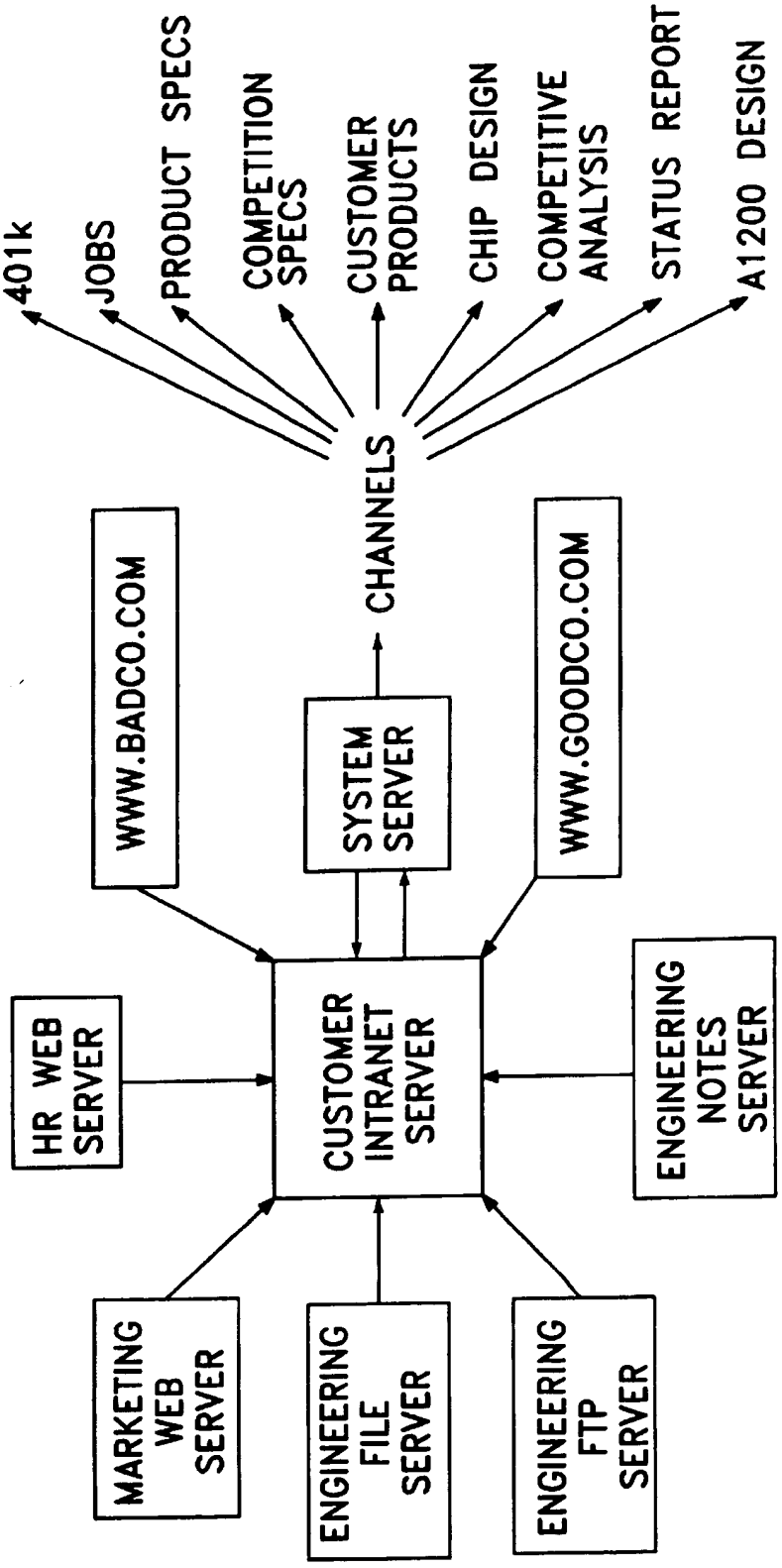

401k
JOBS
PRODUCT SPECS
COMPETITION SPECS
CUSTOMER PRODUCTS
CHIP DESIGN
COMPETITIVE ANALYSIS
STATUS REPORT
A1200 DESIGN
CHANNELS
WWW.BADCO.COM
SYSTEM SERVER
WWW.GOODCO.COM
HR WEB SERVER
CUSTOMER INTRANET SERVER
ENGINEERING NOTES SERVER
MARKETING WEB SERVER
ENGINEERING FILE SERVER
ENGINEERING FTP SERVER

FIG. 2

METHOD AND SYSTEM FOR PROVIDING ACCESS FOR CATEGORIZED INFORMATION FROM ONLINE INTERNET AND INTRANET SOURCES

FIELD OF THE INVENTION

This invention relates to the field of electronic content provision. More specifically, it relates to gathering related content from internet and intranet sources and providing access to same in response to user requests.

BACKGROUND OF THE INVENTION

A huge quantity of information is being continuously created and made available via electronic communications systems. There is so much information that it is simply not possible for an individual person to read it all. On the other hand, it is imperative that certain items of information reach certain people. Much of the electronically-provided news information ages rapidly, such that it loses its relevancy in a matter of days, or even a matter of hours (e.g., stock market information). Each person has different needs for information, and requires access to a different subset of the available information. In light of the foregoing, there is clearly a need for a system and method for rapidly accessing categorized electronic information.

One difficulty in providing the information is that the information is being created in many different places. News articles about events in the world or business community, and articles written for newspapers, magazines and journals, can generally be obtained through various content providers, who frequently aggregate the information from a number of sources into single continuous electronic streams. No content provider today, however, provides access to all available information, so there is a trade-off between full access and complexity. Moreover, an individual user is forced to subscribe to a host of services in order to obtain the information which is generated from different sources, in different countries, and in various languages. Subscribing to many services to some extent negates the benefits realized by the content aggregation by providers, since the user must then filter through multiple copies of the same documents.

Internally, organizations face similar issues. Memos, announcements, documents of various kinds, and intranet web content are created at multiple locations throughout an organization, yet are generally not readily available to all members of the organization. Therefore, the process of collecting the information from all points of origins is a key issue, along with categorization and controlled dissemination of that information.

Another aspect of the problem is the actual matching process, comprising matching the collected and categorized content with an individual user's interests. For matching to work, an individual user must be able to express a diverse set of interests, not just one interest. A language of some kind is necessary to provide a medium for this expression of the user's interest. Further, a system is needed to capture the language and apply it to the items of information. Moreover, the language must embody some kind of high level semantic knowledge, since past word-search-based systems have fallen short of a satisfactory solution. The ability to express, capture and apply a person's interests or needs is a critical feature of the problem.

Finally, there is a need to deliver the information to people who have expressed an interest. The primary requisites for delivery are making sure that access to the information is convenient, even in dynamic situations, and making sure that delivery can occur quickly once the information becomes available. Moreover, people are increasingly mobile and have varied styles of working and of accessing and processing information. An effective delivery system will therefore require that the means of access be ubiquitous, that multiple means of access be available, and that delays in making the information available be minimized.

It is therefore an objective of the present invention to provide a system for gathering, categorizing, and delivering electronic content to users in response to user requests.

It is another objective of the invention to provide a system and method for gathering content from both inside (i.e., intranet) and outside (i.e., internet) sources and categorizing same for provision in response to customized user requests.

Yet another objective of the present invention is to provide a customer with the ability to embed user interest and delivery mechanisms into customer applications.

SUMMARY OF THE INVENTION

These and other objectives are realized by the present invention which provides a system for collecting and categorizing metadata about content provided via the internet or intranet; for gathering user interest information and creating a user profile for matching to collected and categorized content information; and for matching and delivering categorized information tailored to customized user profiles.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention will now be further detailed with specific reference to the appended figures wherein:

FIG. **1** provides a schematic illustration of an implementation of the present invention.

FIG. **2** provides a schematic illustration of the intranet side of one embodiment of the inventive system.

DESCRIPTION OF THE PREFERRED EMBODIMENT

FIG. 1 provides a schematic overview of one implementation of the present invention. The implementation can be viewed as having two sides, an external side comprising the sources, content providers and System Server, and an internal side comprising the customer's site (including at least one server for the customer's intranet), internal sources and end user sites. As detailed therein, sources **11** provide electronic content (e.g., articles) on-line to content providers **12**. The System Server, **10**, gathers electronic content from content providers, as well as directly from sources, if necessary. At the System Server, the electronic content is categorized, with duplicate copies eliminated, and is stored in so-called "channels" of information. Each channel represents a particular category or group of categories of related information. The categorization of document content is generally done without reference to known user profiles or prejudices, although the categorization can be influenced by known or expected user query categories.

On the so-called "internal" side of the inventive system, the Customer Intranet Server **14**, is in communication with not only the outside System Server, but also with internal sources **15** and at least one end user **16**. The end users may be employees of the customer or clients of the customer who have contracted or otherwise arranged for receipt of information which has been accumulated, categorized, and disseminated from the Customer Intranet Server site **14**.

An end user **16** will specify the areas of interest for which that end user wishes to obtain electronic information. Unlike

prior art systems which allowed only minimal user query input, often limited to single word entries for simple word searching, the present system assembles a complex user query including the specification of multiple disparate topics of interest. The user profile is created by system components which are located at the Customer Intranet Server **14**. "Creation" of the user profile involves not only the extension of user-input language, but also the elimination of non-critical language, inclusion of semantic knowledge, and cross-relating of user interest topics. Query development is further detailed below.

Once the user profile has been developed, it is stored at the Customer Intranet Server for matching to assembled and categorized content. The system can be programmed to conduct on-going matching (i.e., checking every new document entry for a match with the user profile), periodic matching (e.g., every 12 hours), or matching only upon user prompting (e.g., only when a user connects to the Customer Intranet Server and asks for an update).

Continual or periodic categorization of external electronic content is the task of the system components which can be preferably located at the System Server **10**. The System Server receives input from the content providers **12**, as well as possibly from the internal sources **15** via the Customer Intranet Server **14**. Receipt of input from both external and internal sources can be a passive process, whereby the documents are continuously or periodically supplied to the System Server, or an active process, whereby system crawler components seek out the documents via word searching, site mapping, etc.

The inventive system preferably includes provision to the customer site of at least one internal crawler which will provide a totally automated way to bring their entire distributed network of resources into the system. The crawlers crawl through a customer's internal network and retrieve documents from various sources, distinguished by the technologies which were used to store the information.

Documents from the internal sources are assembled and categorized at the Customer Intranet Server where a Channel Map is created containing a list of web servers, directories and other targets which have been or are to be crawled. A Channel Map also can be constucted at the System Server as well. Each entry in the Channel Map may include a list of channels in which web pages and documents from the respective server and directory are to appear. Table 1 provides a sample Channel Map for a fictitious semiconductor manufacturer:

TYPE	SERVER	DIRECTORY	CHANNELS
Web	HR	/publish/benefits/401k	401k
Web	HR	/publish/jobopenings	Jobs
Web	Marketing	/publish/product/specs	Product Specs
Web	www.badco.com	/pub/productspecs	Competition Specs
Web	www.goodco.com	/pub/products/electrnic	Customer Products
PCFile	engineering	/projects/chipdesigns	Chip Designs
PCFile	marketing	/reports/companalysis	Competitive Anly.
FTP	engineering	/projects/status	Status Reports
Notes	engineering	/specs/chipspeed	A1200 Design

FIG. **2** provides a schematic illustration of the sources accessible to the Customer Intranet Server of the fictitious company, directly or through the System Server, and the channels that result from receiving or crawling those sources. Information gathered from external sources will also be mapped to the established channels, so that an end user can readily access all relevant information in a category or channel as the result of a single query.

While some amount of categorization may be straightforward, such as those above-noted examples wherein any information obtained from a certain source will necessarily be provided on a given channel (i.e., with sites or site directories being mapped to the channels), the bulk of document categorization requires intensive analysis of the document contents. In addition to the crawlers which automatically funnel documents obtained from certain sources into pre-established channels, there are two other primary means by which documents are categorized. The first, and most rudimentary, is categorization by manual user interface, whereby a system administrator (or even document author) identifies the document to be loaded into the server and identifies the channels in which the document is to appear. The second, more complex, means is automatic categorization by content filtering, which is conducted by system components located at either at the Customer Intranet Server, or at the System Server **10**, the details of which are further provided below and in co-pending applications, Ser. No. 08/979,248, entitled "Method and System for Electronic Document Content or Query Content Filtering", and Ser. No. 08/980,075, entitled "Content Filtering for Electronic Documents Generated in Multiple Foreign Languages", which are assigned to the present assignee, and are being filed on even date herewith. Such automatic categorization can also be utilized at the Customer Intranet Server for the purpose of categorizing internal documents into channels, which may match or be unique from the channels provided by the System Server. Such channel definitions can be applied as well to documents received from the System Server to fill customer-defined channels with news or other external documents. After query processing and document content categorization, it is desirable to analyze the categories to ascertain if other relationships exist among the categories, which relationships themselves may be identified as new categories or channels. Such category processing is the subject of co-pending patent application, Ser. No. 08/978,712, entitled "Category Processing of Query Topics and Electronic Document Content Topics", which is assigned to the present assignee, and is being filed on even date herewith.

Once documents from both the internal and external sources have been categorized/assigned channels, both the documents and the assigned channels are stored in a local database at the Customer Intranet Server or associated customer location. Inventive components at the Customer Intranet Server match the channels assigned to each of the incoming documents with the user's interests as found in the user profile. Each document is then made available for access by, or is sent to, the user whose interests it matches.

The System Server's above-noted functions may be provided as part of a customer intranet, wholly outside of the customer domain, or divided in function between the two locations. In the "outside" example, all document collection and categorization would be done at the System Server as a service of the provider. Documents found on the external internet, as well as those which may be supplied from the customer's own intranet and/or databases, would be analyzed and categorized at the provider location. In the instance where the customer wishes to additionally be a provider to end users, two alternative scenarios are possible. Under the first scenario, an outside provider would still assemble and categorize documents from outside sources and make them available at the customer's server. The customer's server would also be adapted to perform assem-

bly and categorization of "in-house" documents, merging of the in-house assemblage with the categorized documents from outside sources, matching the resultant merged documents to user request profiles, and disseminating the matching results to the user. The second alternative implementation would locate all categorization functionality at the customer location. In all three implementations, the customer location would retain the capability for receipt of user request input, creation and storage of the user profile, matching of the user profile to the categories or channels into which the documents are placed, and provision of the matched documents for end user review.

The customer site is provided with the capability for building applications to create a series of different user interfaces with different interaction means, different restrictions for user access (e.g., providing some users access to only documents from outside sources, while others would have access to both externally-obtained and internally-generated documents), and different levels of query and content complexity.

For the following detailed description of one refinement of one of the processing "stages," including user query analysis and profile creation, document categorization, and matching, it is to be noted that the same types of analyses can frequently be applied at each stage. For example, finding relationships between two seemingly disparate user query subject categories can parallel the effort to identify commonality of subject matter from two input documents, as well as a subsequent effort to match the profile to a category/channel. Therefore, where appropriate, the ensuing processes will reference one, two or all three of the stages of profile or query analysis, document content categorization, and matching stages.

Users of the system initially specify which topics are of interest. This specification may take the form of a simple subscription to pre-defined user interest categories, a modifiable subscription whereby the user may add to or otherwise edit the pre-defined categories, a completely user-customized set of queries, or a combination of any of the foregoing. Each query represents a topic, and can identify a channel and additionally contain boolean, fuzzy, proximity and/or hierarchical operators. A set of topics preferred by a user is known as a user profile. The present method reduces each query to one or more vector entries with the entry's index into the vector corresponding to a hash of the query's textual expression of the importance of that query to the overall topic/profile.

Automatic query processing, as well as document content categorization, is optimized in the present invention by first tokenizing the content thereof. In such a tokenization process, all the word/phrases are first identified as units, then stemmed (e.g., "advanc" will represent "advancement", "advancing", "advance"). After all stop words and phrases are filtered out by processes which are the subject of one of the aforementioned co-pending applications, only a few of the original word/phrases are left. These surviving words/phrases are called tokens. A query can be either a single token (word or phrase) or a combination of tokens which includes boolean, fuzzy, proximity and/or hierarchical operators. Token IDs are assigned to each query item, with the tokens usually being just the stems of the original words, or made-up labels which correspond to phrases. The stems or made-up labels are referred to as "terms". Terms are strings, and since the system must handle quite a few thousand terms, the total memory which can be consumed by terms can take up a significant amount of computer memory. Therefore, a hash function is provided to assign unique token IDs to the terms (which may also consist of expressions containing words and phrases combined with a variety of query operations). The term strings are replaced by 32 bit integers. The 32 bit integers are then used to represent positions in a vector. Each position in a vector representing a document contains a count of the number of times the term strings appeared in the query.

Textual content can likewise be mapped to vectors using the same procedures as were used for the query topics, above. Each channel definition is also represented by a vector reflecting the term strings found in documents belonging to that channel. A "reverse dictionary" can be maintained which comprises a lexicon with token IDs as the keys and the words, phrases and/or queries as the values. However, if the need is to mark the document with categories and to retrieve based on the specific tokens matched, without a need to "translate" the assigned tokens back to categories, a lexicon will not be needed.

Clearly, with the present system and method, when comparisons are being made, comparisons of 32 bit integers will be significantly faster than the prior art string comparisons. Query vectors are compared to (e.g., each vector entry is serially processed against) the channel definition vectors and/or the document vectors. Once the vectors are normalized, classification and matching are reduced to vector processing, such as the processing detailed in the aforementioned co-pending patent applications.

The invention has been described with reference to several specific embodiments. One having skill in the relevant art will recognize that modifications may be made without departing from the spirit and scope of the invention as set forth in the appended claims.

Having thus described our invention, what we claim as new and desire to secure by Letters Patent is:

1. A method for a provider computer system to provide user access to electronic documents comprising the steps of:

obtaining at least one user profile of user interests by receiving input from said user and determining query categories for said user profile from said input;

gathering a plurality of first electronic documents from at least one electronic content source external to said provider computer location;

receiving a plurality of second electronic documents from at least one electronic content source local to said provider computer system;

categorizing said plurality of first electronic documents by topic categories;

organizing said plurality of second electronic documents into topic categories;

storing said topic categories and said first and second electronic documents at said provider computer location;

matching each of said at least one user profiles to said topic categories; and

retrieving said stored first and second electronic documents in topic categories matched to said at least one user profile;

wherein said determining comprises the steps of:

organizing said user input into a plurality of query items;

assigning a token ID to each of said plurality of query items; and

hashing each said token ID into a query vector.

2. The method of claim 1 further comprising correlating said query categories to known topic categories.

3. The method of claim 1 wherein said gathering of said plurality of first electronic documents comprises receiving packets of documents from external content providers.

4. The method of claim 1 wherein said gathering of said plurality of first electronic documents comprises crawling sites of said external content sources for documents.

5. The method of claim 1 wherein said receiving said plurality of second electronic documents comprises receiving packets of documents from said at least one local content source.

6. The method of claim 1 wherein said gathering of said plurality of second electronic documents comprises crawling sites of said at least one local content source for documents.

7. The method of claim 1 wherein said gathering and said receiving is done continually.

8. The method of claim 1 wherein said gathering and said receiving is done periodically.

9. The method of claim 1 wherein said gathering and said receiving is done in response to a user request.

10. The method of claim 1 wherein said organizing of said plurality of second electronic documents comprises mapping each of said second documents into topic categories based upon the location of said local electronic content source.

11. The method of claim 1 wherein said organizing of said plurality of second electronic documents comprises assigning each of said second documents to topic categories based upon manual assignment of topic categories to said documents.

12. The method of claim 1 wherein said organizing of said plurality of second electronic documents comprises assigning each of said second documents to topic categories based upon automatic content filtering.

13. The method of claim 1 wherein said categorizing of said plurality of first electronic documents comprises assigning each of said second documents to topic categories based upon manual assignment of topic categories to said documents.

14. The method of claim 1 wherein said categorizing of said plurality of first electronic documents comprises mapping each of said first documents into topic categories based upon said external electronic content source.

15. The method of claim 1 wherein said categorizing of said plurality of first electronic documents comprises assigning each of said first documents to topic categories based upon automatic content filtering.

16. A method of claim 1 wherein said first electronic documents are categorized into first topic categories, said second electronic documents are categorized into second topic categories and wherein said method further comprises merging said first and said second topic categories.

17. The method of claim 1 further comprising refining said merged topic categories based on said at least one user profile.

18. A method for a provider computer system to provide user access to electronic documents comprising the steps of:

obtaining at least one user profile of user interests;

gathering a plurality of first electronic documents from at least one electronic content source external to said provider computer location;

receiving a plurality of second electronic documents from at least one electronic content source local to said provider computer system;

categorizing said plurality of first electronic documents by topic categories;

organizing said plurality of second electronic documents into topic categories;

storing said topic categories and said first and second electronic documents at said provider computer location;

matching each of said at least one user profiles to said topic categories; and

retrieving said stored first and second electronic documents in topic categories matched to said at least one user profile;

wherein said categorizing comprises the steps of:

organizing content from said first electronic documents into a plurality of items;

assigning a token ID to each of said plurality of items; and

hashing each said token ID into a document vector; and

said organizing comprises the steps of:

organizing content from said second electronic documents into a plurality of items;

assigning a token ID to each of said plurality of items; and

hashing each said token ID into a document vector.

19. The method of claim 1 wherein said categorizing comprises the steps of:

organizing content from said first electronic documents into a plurality of items;

assigning a token ID to each of said plurality of items; and

hashing each said token ID into a document vector; and

said organizing comprises the steps of:

organizing content from said second electronic documents into a plurality of items;

assigning a token ID to each of said plurality of items; and

hashing each said token ID into a document vector.

20. The method of claim 19 further comprising assigning a topic token ID to each of said topic categories; and hashing each of said topic token IDs into a category vector.

21. The method of claim 20 wherein said matching comprises vector processing of query categories against said category vectors.

22. The method of claim 21 wherein said matching further comprises vector processing of query categories against document vectors.

* * * * *